Intro

RoadMap

TODO

Basics and Linear System Of Equations (LSEQ) Intro

If not mentioned, consider dimension of matrix A to be m x n

Matrix Multiplication

$c_{jk} = a_j b_k$, where a_j is the jth row vector of A and b_k is the kth column vector of B

Using this matrix multiplication can be computed parallelly (product is computed column wise)

$$AB = A[B_1 B_2 \dots B_p] = [AB_1 AB_2 \dots AB_p] = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} B = \begin{bmatrix} A_1 B \\ \vdots \\ A_m B \end{bmatrix}$$

- Matrix multiplication is associative.

Motivation of Multiplication by Linear Transformation

Suppose $x_1 x_2$-coordinate system is related to a $y_1 y_2$-coordinate system as

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} x_1 + a_{12} x_2 \\ a_{21} x_1 + a_{22} x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = Ax$$

Now suppose further that the $x_1 x_2$-system is related to a $w_1 w_2$-system by another linear Transposition, say,

$x = Bw$

Then y is related to w (which if you go by substitution of variables) as $y = ABw$.

Transposition properties

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(cA)^T = c(A)^T$

- $(AB)^T = B^T A^T$ –> *This prop. can be extended*

Complex Conjugate

$\bar{A} = [\bar{a_{ij}}]$

- $\overline{kA} = \bar{k}\bar{A}$
- $\overline{AB} = \bar{A}\bar{B}$
- $\overline{A^T} = \bar{A}^T$ (Obvious), and $\overline{A^T}$ is also denoted simply as A^θ
- $(AB)^\theta = B^\theta A^\theta$
- $(kA)^\theta = \bar{k}A^\theta$

Special Square Matrices

All these are defined solely for **square** matrices.

- Symmetric: $A^T = A$
- Skew-Symmetric: $A^T = -A$ so diagonal elements must be zero (diagonal is defined only for square matrices).
- Upper Triangular: $a_{ij} = 0$ if $i > j$
- Lower Triangular: $a_{ij} = 0$ if $j > i$
- Hermitian: $\overline{A^T} = A$, *i.e., equal to conjugate transpose* (Thus elements in diagonal must be real)
- Skew-Hermitian: $\overline{A^T} = -A$
- Scalar: $S = cI$
- Orthogonal: $A^T A = I$ (thus $|A| \neq 0 (= \pm 1)$ and $A^{-1} = A^T$)
- Unitary: $A^\theta A = I$ (thus $|A| \neq 0 (= e^{it}$ as $\overline{|A|}|A| = ||A||^2 = 1$ and $A^{-1} = A^\theta$)
- Involutory: $A^2 = I$
- Nilpotent: Square matrix A such that there exist some positive integer n for which $A^n = O$. Such a smallest positive integer is called an index.
- Idempotent: $A^2 = A$
- Singular: $|A| = 0$ o/w non singular

Trace

- $tr(AB) = tr(BA)$ (easy to prove)

Determinant

- $(-1)^{i+j} M_{ij} = C_{ij}$

- Determinant (which is defined only for square matrix) remains the same when expanded along any row or column.

- $|A| = |A^T|$

- $|AB| = |A||B|$ So if $|AB| = 0 \to |A| = 0$ or $|B| = 0$

  $AB = O$ doesn't imply $A = O$ or $B = O$ but that atleast one of them is singular.

- Swapping any 2 rows or any 2 columns multiplies the determinant by -1

- $\begin{bmatrix} ka_1 & b_1 & c_1 \\ ka_2 & b_2 & c_2 \\ ka_3 & b_3 & c_3 \end{bmatrix} = k \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$

- $\begin{bmatrix} ka_1 & kb_1 & kc_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = k \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$

- $|kA_n| = k^n |A_n|$

- If 2 rows (or 2 columns) are identical, then determinant is zero.

- Sum of the product of elements of any row (or column) with the cofactors of the corresponding elements of any other row (or column) is zero. Can be used to prove $A\,adj(A) = |A|I$. (*Note*: Adjoint $= C^T$ where C is the cofactor matrix)

- Adding a constant multiple of a row (or column) to another row (or colum) doesn't change the determinant.

- $|\bar{A}| = \overline{|A|}$ (Can be easily shown by induction or even otherwise as $\overline{a + cb} = \bar{a} + \bar{b}\bar{c}$)

- $|A^\theta| = \overline{|A|}$ ($|A^\theta| = |\bar{A}^T| = \overline{|A^T|} = \overline{|A|}$)

- Determinant of Hermitian is always a real number. To show: $|A| = \overline{|A|}$ i.e., to show $|A^\theta| = \overline{|A|}$, which is true.

- Similarly determinant of a skew hermitian matrix is always imaginary.

- Determinant of skew symmetric matrix of odd order is 0.

Adjoint

$(adj(A))A = A(adj(A)) = |A|I \to A^{-1} = adj(A)/|A|$, as evident inverse is defined only for square matrices and exists iff $|A| \neq 0$.

- $adj(X) = |X|X^{-1}$

- $adj(AB) = adj(B)adj(A)$ (follows from above)

- $(A^{\theta})^{-1} = (A^{-1})^{\theta}$

  $A^{\theta}B = I \rightarrow B^{\theta}A = I$

  $\rightarrow B^{\theta} = A^{-1} \rightarrow B = (A^{-1})^{\theta}$
- $(A^T)^{-1} = (A^{-1})^T$ (easy)
- $adj(kA) = k^{n-1}adj(A)$ (as cofactor will consist of determinant of order $n-1$)
- $adj(A^T) = adj(A)^T$ (proof omitted)
- If A is symmetric then $adj(A)$ is symmetric.
- $|A^{-1}| = |A|^{-1}$ (easy, thus from this it follow $|adj(A)| = |C| = |A|^{n-1}$)

Linear System Of Equations Intro

A linear system of m equations in n unknowns $x_1, \ldots, x_n$ is a set of equations of the form

$a_{11}x_1 + \cdots + a_{1n}x_n = b_1$

$a_{21}x_1 + \cdots + a_{2n}x_n = b_2$

$a_{m1}x_1 + \cdots + a_{mn}x_n = b_m$

The system is called linear because each variable x_j appears in the first power only. If all the b_j are zero, then this is called a homogeneous system o/w nonhomogeneous.

For a homogeneous system we always have trivial soln: $x_i = 0$

This system can be represented as $Ax = b$, a_{ij} are called coefficients and thus A is called coefficient matrix.

$$\tilde{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} & | & b_1 \\ . & \cdots & . & | & . \\ . & \cdots & . & | & . \\ a_{m1} & \cdots & a_{mn} & | & b_m \end{bmatrix} \rightarrow \text{is called Augmented Matrix, it determines}$$
the system completely

Elementary Row Operations

- Interchange of two rows
- Addition of a constant multiple of one row to another row
- Multiplication of a row by a **nonzero** constant c

Similarly we have **Elementary Column Operations**

We now call a linear system S_1 row-equivalent to a linear system S_2 if S_1 can be obtained from S_2 by (finitely many!) row operations (observe that it is an equivalence relation).

A system is called consistent if it has at least one solution (thus, one solution or infinitely many solutions), but inconsistent if it has no solutions at all.

Elementary Matrix

Matrix which differs from Identity Matrix by one single elementary row operation. Left multiplication (pre-multiplication) by an elementary matrix represents elementary row operation, while right multiplication (post - multiplication) represents elementary column operations.

- E_{ij}, $E_i(k)$ and $E_{ij}(k)$, denotes elementary matrix obtained by:
    - swapping rows i and j
    - multiplying ith row with non zero k
    - adding to the ith row the constant multiple (k) of the jth row, resp.
    - Clearly $E_{ij}^{-1} = E_{ij}$, $E_i(k)^{-1} = E_i(\frac{1}{k})$ and $E_{ij}(k)^{-1} = E_{ij}(-k)$ thus inverse of an elementary matrix is also an elementary matrix.
- $|E_{ij}| = -1$
- $|E_i(k)| = k$
- $|E_{ij}(k)| = 1$
- Thus elementary matrix is always non singular

Row Echelon Form

For each row in a matrix, if the row does not consist of only zeros, then the leftmost nonzero entry is called the leading coefficient (or pivot) of that row. So if two leading coefficients are in the same column, then a row operation of type 3 could be used to make one of those coefficients zero. Then by using the row swapping operation, one can always order the rows so that for every non-zero row, the leading coefficient is to the right of the leading coefficient of the row above. If this is the case, then matrix is said to be in row echelon form.

Reduced Row Echelon Form

A matrix is in reduced row echelon form (also called row canonical form, row reduced echelon form) if it satisfies the following conditions:

- It is in row echelon form.
- The leading entry in each nonzero row is a 1 (called a leading 1).
- Each column containing a leading 1 has zeros everywhere else (i.e. also in above).

Side Problems

- If P and Q are non singular matrices then show that $\begin{bmatrix} P & O \\ O & Q \end{bmatrix}^{-1} = \begin{bmatrix} P^{-1} & O \\ O & Q^{-1} \end{bmatrix}$

  Let inverse be $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

  $\rightarrow \begin{bmatrix} PA & PB \\ QC & QD \end{bmatrix} = \begin{bmatrix} I & O \\ O & I \end{bmatrix}$ ∎

Vector Spaces, LSEQ and Rank of a Matrix

Field

Field has two operations, called addition and multiplication; it is an abelian group under the addition, with 0 as additive identity; the nonzero elements are an abelian group under the multiplication, with 1 as multiplicative identity; the multiplication is distributive over the addition.

Subfield

Let F be field and $K \subseteq F$. If K is a field wrt same binary operations in F then K is called subfield of F.

Example:

- I is not a subfield of Q
- R is a subfield of C

Internal Composition

Let A be any set. If $a * b \in A \forall a, b \in A$ then $*$ is said to be internal composition on A

External Composition

Let V and F be any two sets. If $a \circ v \in V \forall a \in F, v \in V$ then $\circ$ is said to be an external composition in V over F.

Vector Spaces

A set V is said to be a vector space over a field F (denoted as $V(F)$) if V is an Abelian group under addition (denoted by $+$) and, if for each $a \in F$ and $v \in V$, there is an element av in V such that the following conditions hold for all a, b in F and all u, v in V.

1. $a(v + u) = av + au$
2. $(a + b)v = av + bv$
3. $a(bv) = (ab)v$
4. $1v = v$

- $k0 = 0$ (easy)

- $0u = 0$ (easy)

- if $ku = 0 \rightarrow k = 0$ or $u = 0$ (Suppose $k \neq 0 \rightarrow \exists k^{-1}$ s.t. $kk^{-1} = 1$ $\rightarrow k^{-1}(ku) = 0 \rightarrow 1u = 0 \rightarrow u = 0$)

- $k(-u) = -ku$ (As $k(u - u) = 0 = ku + k(-u) \rightarrow -ku = k(-u)$))

- $(-k)u = -ku$ (Proof similar as above)

- $u + u = 2u$ (As $1u + 1u = (1 + 1)u = 2u$)

- Thus, an internal composition in V is called vector addition and external composition in V over the field F is called scalar multiplication.

- If $F = \mathbb{R}, \mathbb{Q}, \mathbb{C}, ...$ then V is resp. called Real vector space, Rational vector space, Complex vector space, ...

- A field F is a vector space over any subfield K of F.

- Let K be an arbitrary field, K^n is used to denote the set of all n tuples of elements in K. Here K^n is a vector space over K using the following operations:

  1. Vector Addition: $(a_1, a_2, ... , a_n) + (b_1, b_2, ... , b_n) = (a_1 + b_1, a_2 + b_2, ... , a_n + b_n)$
  2. Scalar Multiplication: $k(a_1, a_2, ... , a_n) = (ka_1, ka_2, ... , ka_n)$

- Let $P_n(t)$ denote the set of all polynomials $p(t)$ over a field K, where the degree of $p(t)$ is less than or equal to n; that is, $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_s t^s$ where $s \leq n$. Then $P_n(t)$ is a vector space over K with respect to the usual operations of addition of polynomials and of multiplication of a polynomial by a constant. We include the zero polynomial 0 as an element of $P_n(t)$, even though its degree is undefined.

- **Function Space:** Let X be a nonempty set and let K be an arbitrary field. Let $F(X)$ denote the set of all functions of X into K. [Note that $F(X)$ is nonempty, because X is nonempty.] Then $F(X)$ is a vector space over K with respect to the following operations:

  1. Vector Addition: The sum of two functions f and g in $F(X)$ is the function $f + g$ in $F(X)$ defined by $(f + g)(x) = f(x) + g(x) \quad \forall x \in X$
  2. Scalar Multiplication: The product of a scalar $k \in K$ and a function f in $F(X)$ is the function kf in $F(X)$ defined by $(kf)(x) = kf(x) \quad \forall x \in X$

Spanning Sets

Let V be a vector space over K. Vectors $u_1, u_2, \ldots, u_m$ in V are said to span V or to form a spanning set of V if every v in V is a linear combination of the vectors $u_1, \ldots, u_m$ that is, if there exist scalars $a_1, a_2, \ldots, a_m$ in K such that
$v = a_1 u_1 + a_2 u_2 + \cdots + a_m u_m$

Remark 1: Suppose $u_1, u_2, \ldots, u_m$ span V and suppose u_k is a linear combination of some of the other u' s. Then the u' s without u_k also span V.

Remark 2: Suppose $u_1, u_2, \ldots, u_m$ span V and suppose one of the u' 's the zero vector. Then the u' 's without the zero vector also span V .

Subspaces

Let V be a vector space over a field K and let W be a subset of V. Then W is a subspace of V if W is itself a vector space over K with respect to the operations of vector addition and scalar multiplication on V.

Theorem: Suppose W is a subset of a vector space V. Then W is a subspace of V if the following two conditions hold:

1. The zero vector 0 belongs to W.
2. For every $u, v \in W, k \in K$: (i) The sum $u+v \in W$. (ii) The multiple $ku \in W$

Both properties may be combined into the following equivalent single statement:

For every $u, v \in W, a, b \in K$, the linear combination $au + bv \in W$

- Now let V be any vector space. Then V automatically contains two subspaces: the set $\{0\}$ consisting of the zero vector alone and the whole space V itself. These are sometimes called the trivial subspaces of V.

- Let V be the vector space of real-valued functions. Then the collection W_1 of continuous functions and the collection W_2 of differentiable functions are subspaces of V.

Intersection of Subspaces

Let U and W be subspaces of a vector space V. We show that the intersection $U \cap W$ is also a subspace of V. Clearly, $0 \in U$ and $0 \in W$, because U and W are subspaces; whence $0 \in U \cap W$. Now suppose u and v belong to the intersection $U \cap W$ now because U and W are subspaces, for any scalars $a, b \in K$, $au + bv \in U$ and $au + bv \in W$

Thus, $au + bv \in U \cap W$. Therefore, $U \cap W$ is a subspace of V.

The above result generalizes as follows.

Theorem: The intersection of any number of subspaces of a vector space V is a subspace of V.

Similarly,

Theorem: The solution set W of a homogeneous system $AX = 0$ in n unknowns is a subspace of K^n.

Linear Spans

Suppose $u_1, u_2, \ldots, u_m$ are any vectors in a vector space V. The collection of all linear combinations of them is denoted by $span(u_1, u_2, \ldots, u_m)$ or $span(u_i)$ and is called linear span of them.

It is easy to see that $span(u_i)$ is a subspace of V.

More generally, for any subset S of V, $span(S)$ consists of all linear combinations of vectors in S or, when $S = \phi$, $span(S) = \{0\}$.

Thus, in particular, S is a spanning set of $span(S)$.

Theorem: Let S be a subset of a vector space V.

1. Then $span(S)$ is a subspace of V that contains S. (easy to prove)
2. If W is a subspace of V containing S, then $span(S) \subseteq W$ (i.e. $span(S)$ is the 'smallest' subspace of V containing S). (easy to prove)

- Consider the vector space $V = R^3$
  - Let u be any non zero vector in R^3. Then $span(u)$ consists of all scalar multiples of u. Geometrically, $span(u)$ is the line through the origin O and the endpoint of u
  - Let u and v be vectors in R^3 that are not multiples of each other. Then $span(u, v)$ is the plane through the origin O and the endpoints of u and v

Row space of a matrix

Let $A = [a_{ij}]$ be an arbitrary $m \times n$ matrix over a field K. The rows of A,

$$R_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \quad R_2 = (a_{21}, a_{22}, \ldots, a_{2n}), \quad \ldots, \quad R_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$$

may be viewed as vectors in K^n; hence, they span a subspace of K^n called the row space of A and denoted by $rowsp(A)$. That is, $rowsp(A) = span(R_1, R_2, \ldots, R_m)$ Analgously we have column span, $colsp(A) = rowsp(A^T)$

- Suppose M is the matrix obtained by applying elementary row operations on a matrix A. Then each row of M is a row of A or a linear combination of rows of A. Hence, the row space of M is contained in the row space of A. On the other hand, we can apply the inverse elementary row operation on M to obtain A; hence, the row space of A is contained in the row space of M. Accordingly, A and M have the same row space. Thus, Row equivalent matrices have the same row space.

  In the context of lseq, this would mean that row equivalent matrices have the same set of solution because they span same set of equations. (Note well that we are dealing with row operations, No column operations on the augmented matrix are permitted in this context because they would generally alter the soln set)

Theorem: Suppose $A = [a_{ij}]$ and $B = [b_{ij}]$ are row equivalent echelon matrices with respective pivot entries $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ and $b_{1k_1}, b_{2k_2}, \ldots, b_{sk_s}$ Then A and B have the same number of nonzero rows—that is, $r = s$—and their pivot entries are in the same positions—that is, $j_1 = k_1, j_2 = k_2, \ldots, j_r = k_r$. (Easy as o/w row space is clearly different)

Theorem: Suppose A and B are row canonical matrices. Then A and B have the same row space if and only if they have the same nonzero rows.

Proof: Consider ith non zero row of A. Consider vectors belonging to its row space having only that pivot element set to 1 and rest all pivot elements set to 0. That means only this row is multiplied by 1 and rest all rows are multiplied by 0. So to get this vector in B, we would have to keep only this ith row multiplied by 1 and rest all zero. Thus the rows are same.

Corollary: Every matrix A is row equivalent to a unique matrix in row canonical form.

- Consider the following two sets of vectors in R^4:

  $u_1 = (1, 2, -1, 3), \quad u_2 = (2, 4, 1, -2), \quad u_3 = (3, 6, 3, -7)$

  $w_1 = (1, 2, -4, 11), \quad w_2 = (2, 4, -5, 14)$

  Let $U = span(u_i)$ and $W = span(w_i)$. There are two ways to show that $U = W$.

  1. Show that each u_i is a linear combination of w_1 and w_2, and show that each w_i is a linear combination of u_1, u_2, u_3. Observe that we have to show that six systems of linear equations are consistent.

2. Form the matrix A whose rows are u_1, u_2, u_3 and row reduce A to row canonical form, and form the matrix B whose rows are w_1 and w_2 and row reduce B to row canonical form:

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 6 & -16 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 2 & -4 & 11 \\ 2 & 4 & -5 & 14 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -4 & 11 \\ 0 & 0 & 3 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & -\frac{8}{3} \end{bmatrix}$$

Because the nonzero rows of the matrices in row canonical form are identical, the row spaces of A and B are equal. Therefore, $U = W$.

Linear Dependence and Independence

We say that the vectors $v_1, v_2, \ldots, v_m$ in V are linearly dependent if there exist scalars $a_1, a_2, \ldots, a_m$ in K, not all of them 0, such that

$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m = 0$

Otherwise, we say that the vectors are linearly independent.

A set $S = v_1, \ldots, v_m$ of vectors in V is linearly dependent or independent according to whether the vectors $v_1, \ldots, v_m$ are linearly dependent or independent.

An infinite set S of vectors is linearly dependent or independent according to whether there do or do not exist vectors $v_1, v_2, \ldots, v_k$ in S that are linearly dependent.

- Suppose 0 is one of the vectors $v_1, v_2, \ldots, v_m$, say $v_1 = 0$. Then the vectors must be linearly dependent.
- Suppose v is a nonzero vector. Then v, by itself, is linearly independent
- Let V be the vector space of functions from R into R. We show that the functions $f(t) = sint, g(t) = e^t, h(t) = t^2$ are linearly independent. We form the vector (function) equation $xf + yg + zh = 0$, where x, y, z are unknown scalars. This function equation means that, for every value of t, $x \sin t + y e^t + z t^2 = 0$

Thus, in this equation, we choose appropriate values of t to easily get $x = 0, y = 0, z = 0$.

1. Substitute $t = 0$ to obtain $x(0) + y(1) + z(0) = 0$ or $y = 0$
2. Substitute $t = \pi$ to obtain $x(0) + 0 (e^\pi) + z (\pi^2) = 0$ or   or $z = 0$
3. Substitute $t = \pi/2$ to obtain $x(1) + 0 (e^{\pi/2}) + 0 (\pi^2/4) = 0$ or $x = 0$

Thus u, v and w are LI.

Why is linear independence important? Well, if a set of vectors is linearly dependent, then we can get rid of at least one or perhaps more of the vectors

until we get a linearly independent set. This set is then the smallest "truly essential" set with which we can work.

Lemma: Suppose two or more nonzero vectors $v_1, v_2, \ldots, v_m$ are linearly dependent. Then one of the vectors is a linear combination of the preceding vectors; that is, there exists $k > 1$ such that $v_k = c_1 v_1 + c_2 v_2 + \cdots + c_{k-1} v_{k-1}$ (easy to prove)

Basis and Dimension

A set $S = u_1, u_2, \ldots, u_n$ of vectors is a basis of V if it spans V and is linearly independent.

Theorem: Let V be a vector space such that one basis has m elements and another basis has n elements, then, $m = n$.

This theorem is the consequence of the following replacement lemma.

Lemma: Suppose $\{v_1, v_2, \ldots, v_n\}$ spans V, and suppose $\{w_1, w_2, \ldots, w_m\}$ is linearly independent. Then $m \leq n$, and V is spanned by a set of the form $\left\{w_1, w_2, \ldots, w_m, v_{i_1}, v_{i_2}, \ldots, v_{i_{n-m}}\right\}$. Thus, in particular, $n + 1$ or more vectors in V are linearly dependent.

Proof: Because $\{v_i\}$ spans V, we have

$$\{w_1, v_1, \ldots, v_n\} \tag{1}$$

is linearly dependent and also spans V. By previous Lemma, one of the vectors in is a linear combination of the preceding vectors. This vector cannot be w_1, so it must be one of the $v's$, say v_j: Thus we can delete v_j from the spanning set and obtain the spanning set

$$\left\{w_1, v_1, \ldots, v_{j-1}, \quad v_{j+1}, \ldots, v_n\right\} \tag{2}$$

Now we repeat the argument with the vector w_2. That is, because (2) spans V, the set

$$\left\{w_1, w_2, v_1, \ldots, v_{j-1}, \quad v_{j+1}, \ldots, v_n\right\} \tag{3}$$

is linearly dependent and also spans V. Again by previous Lemma, one of the vectors in (3) is a linear combination of the preceding vectors. We emphasize that this vector cannot be w_1 or w_2, because $\{w_1, \ldots, w_m\}$ is independent; hence, it must be one of the $v's$, say v_k. Thus, we can delete v_k from the spanning set (3) and obtain the spanning set $\left\{w_1, w_2, v_1, \ldots, v_{j-1}, \quad v_{j+1}, \ldots, v_{k-1}, \quad v_{k+1}, \ldots, v_n\right\}$ We repeat the argument with w_3, and so forth. At each step, we are able to add one of the $w's$ and delete one of the $v's$ in the spanning set. If $m \leq n$, then we finally obtain a spanning set of the required form: $\left\{w_1, \ldots, w_m, v_{i_1}, \ldots, v_{i_{n-m}}\right\}$

Finally, we show that $m > n$ is not possible. Otherwise, after n of the above steps, we obtain the spanning set $\{w_1, \ldots, w_n\}$. This implies that w_{n+1} is a

linear combination of $w_1, \ldots, w_n$, which contradicts the hypothesis that $\{w_i\}$ is linearly independent.

A vector space V is said to be of finite dimension n or n-dimensional, written $dim(V) = n$ if V has a basis with n elements. The vector space $\{0\}$ is defined to have dimension 0.

Suppose a vector space V does not have a finite basis. Then V is said to be of infinite dimension or to be infinite-dimensional.

Theorem: Let V be a vector space of finite dimension n.

1. Any linearly independent set $S = \{u_1, u_2, \ldots, u_n\}$ with n elements is a basis of V. (Always use this fact if possible to prove that the given n element independent set form the basis)
2. Any spanning set $T = \{v_1, v_2, \ldots, v_n\}$ of V with n elements is a basis of V.

Proof:

1. Suppose it does not $span(V)$ that means there exist a vector v which is not a linear combination of S thus $S \cup \{v\}$ is linearly independent which by above lemma is not possible.
2. Suppose this set is not LI then its dimension is less than n, which we know is not possible as cardinality of basis is always the same.

Similarly it is straight forward to prove the following theorems:

Theorem: that if S spans a vector space V. Then,

1. Any maximum number of linearly independent vectors in S form a basis of V.
2. Suppose one deletes from S every vector that is a linear combination of preceding vectors in S. Then the remaining vectors form a basis of V.

Theorem: Let V be a vector space of finite dimension and let $S = \{u_1, u_2, \ldots, u_r\}$ be a set of linearly independent vectors in V. Then S is part of a basis of V; that is, S may be extended to a basis of V

Theorem: Let W be a subspace of an n-dimensional vector space V. Then $dim(W) \leq n$. In particular, if $dim(W) = n$, then $W = V$. (easy to prove)

Rank of a Matrix

The rank of a matrix A is the maximum number of linearly independent row vectors of A. It is denoted by rank A.

Note further that $rankA = 0$ if and only if $A = 0$. This follows directly from the definition.

Theorem: Row-equivalent matrices have the same rank. So to find rank of the matrix, convert it into row echelon form and count the number of non zero rows.

Theorem: Consider p vectors that each have n components. Then these vectors are linearly independent if the matrix formed, with these vectors as row vectors, has rank p. However, these vectors are linearly dependent if that matrix has rank less than p.

Theorem: The rank r of a matrix A equals the maximum number of linearly independent column vectors of A. Hence A and its transpose A^T have the same rank.

Proof: Let A be an $m \cdot n$ matrix of $rankA = r$. Then by definition of rank, A has r linearly independent rows which we denote by $v_{(1)}, \dots, v_{(r)}$ (regardless of their position in A), and all the rows $a_{(1)}, \dots, a_{(m)}$ of A are linear combinations of those, say,

$a_{(1)} = c_{11}v_{(1)} + c_{12}v_{(2)} + \cdots + c_{1r}v_{(r)}$

$a_{(2)} = c_{21}v_{(1)} + \cdots + c_{2r}v_{(r)}$

$\vdots$

$a_{(m)} = c_{m1}v_{(1)} + \cdots + c_{mr}v_{(r)}$

These are vector equations for rows. To switch to columns, we write in terms of components as n such systems, with $k = 1, \dots, n$,

$a_{1k} = c_{11}v_{1k} + c_{12}v_{2k} + \cdots + c_{1r}v_{rk}$

$\vdots$

$a_{mk} = c_{m1}v_{1k} + c_{m2}v_{2k} + \cdots + c_{mr}v_{rk}$

Treat the LHS as a column vector and RHS as linear combination of r column vectors ($[c_{1i}, c_{2i}, \dots, c_{mi}]$). Now the vector on the left is the kth column vector of A. We see that each of these n columns is a linear combination of the same r columns on the right. Hence A cannot have more linearly independent columns than rows. Now rows of A are columns of the transpose A^T. For A^T our conclusion is that A^T cannot have more linearly independent columns than rows, so that A cannot have more linearly independent rows than columns. Together, the number of linearly independent columns of A must be r, the rank of A. ∎

:::tip Info Therefore dimension of row space (aka row rank) is same as dimension of column space (aka column rank) which is same as rank of the matrix. :::

Theorem: From above two theorems, it follows that if we have p vectors each having n components. If $n < p$, then these vectors are linearly dependent.

Theorem: Column-equivalent (defined analogously) matrices have the same rank. As elementary column operations is row operations on the transpose. Thus it follows that any combination of elementary row and column operations will not change the rank of the matrix and thus can be used to compute rank of it.

:::tip Info

1. Rank of a matrix is same as the order of highest order non zero submatrix. Thus if rank $= n$ then any submatrix of higher order will have determinant $= 0$ and $|A| \neq 0 \iff rank = n$.
2. It is easily observable that if all minors of order k are zero then rank is certainly less than k as minor of higher order will use throughout its computation minors of smaller order.
3. Every non zero matrix have rank ≥ 1 (as we have non zero minor of order 1).
4. From all this it follows that rank of a submatrix is $\leq$ rank of parent. :::

Gauss Elimination

Is basically converting the augmented matrix into row echelon form (using elementary row operations), from $Ax = b$ to $Rx = f$ where R is in row echelon form.

The number of nonzero rows, r, in R is called the rank of R and also the rank of A. Here is the method for determining whether $Ax = b$ has solutions and what they are:

1. No solution. If r is less than m (meaning that R actually has at least one row of all 0s) and at least one of the numbers $f_{r+1}, f_{r+2}, \dots, f_m$ is not zero, then the system $Rx = f$ is inconsistent: No solution is possible.

If the system is consistent (either $r = m$, or $r < m$ and all the numbers $f_{r+1}, f_{r+2}, \dots, f_m$ are zero), then there are solutions.

2. Unique solution. If the system is consistent and $r = n$, there is exactly one solution, which can be found by back substitution.
3. Infinitely many solutions. To obtain any of these solutions, choose values of $x_{r+1}, \dots, x_n$ arbitrarily. Then solve the rth equation for x_r (in terms of those arbitrary values), then the $(r-1)$st equation for x_{r-1}, and so on up the line.

Thus it can be said that the system is consistent if the rank of A and $[A|b]$ is same. :::

Example of basis

- Vector space K^n: $e_1 = (1, 0, 0, 0, \dots, 0, 0), e_2 = (0, 1, 0, 0, \dots, 0, 0), \dots, e_n = (0, 0, 0, 0, \dots, 0, 1)$ These vectors are linearly independent. (For example, they form a matrix in echelon form.) Accordingly, the vectors form a basis of K^n called the usual or standard basis of K^n

- Vector space $M = M_{r,s}$ of all $r \times s$ matrices, let E_{ij} be the matrix with ij-entry 1 and 0's elsewhere. Then all such matrices (in total rs) form a basis of $M_{r,s}$ (As matrix need to be represented as a linear combination of matrices) called the usual or standard basis of $M_{r,s}$. Accordingly, $dim(M_{r,s}) = rs$.

- Vector space $P_n(t)$ of all polynomials of degree $\leq$ n: The set $S = \{1, t, t^2, \dots, t^n\}$ of $n+1$ polynomials is a basis of $P_n(t)$.

- Vector space $P(t)$ of all polynomials: Consider any finite set $S = \{f_1(t), f_2(t), \dots, f_m(t)\}$ of polynomials in $P(t)$, and let m denote the largest of the degrees of the polynomials. Then any polynomial $g(t)$ of degree exceeding m cannot be expressed as a linear combination of the elements of S. Thus, S cannot be a basis of $P(t)$. This means that the dimension of $P(t)$ is infinite. We note that the infinite set $S' = \{1, t, t^2, t^3, \dots\}$ consisting of all the powers of t, spans $P(t)$ and is linearly independent. Accordingly, S' is an infinite basis of $P(t)$.

- The following four vectors in R^4 form a matrix in echelon form: $(1, 1, 1, 1), (0, 1, 1, 1), (0, 0, 1, 1), (0, 0, 0, 1)$ Thus, the vectors are linearly independent, and, because $dim(R^4) = 4$, the four vectors form a basis of R^4.

- Echelon matrix B (whose pivots are circled) is an echelon form of A:

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 2 & 5 & 5 & 6 & 4 & 5 \\ 3 & 7 & 6 & 11 & 6 & 9 \\ 1 & 5 & 10 & 8 & 9 & 9 \\ 2 & 6 & 8 & 11 & 9 & 12 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 2 & 1 & 3 & 1 & 2 \\ 0 & 0 & 3 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore basis of $rowsp(A) = \{(1, 2, 1, 3, 1, 2), \quad (0, 1, 3, 1, 2, 1), \quad (0, 0, 0, 1, 1, 2)\}$

Now to find each column C_k of A that is a linear combination of preceding columns of A.

Let $M_k = [C_1, C_2, \ldots, C_k]$ the submatrix of A consisting of the first k columns of A. Then M_{k-1} and M_k are, respectively, the coefficient matrix and augmented matrix of the vector equation $x_1 C_1 + x_2 C_2 + \cdots + x_{k-1} C_{k-1} = C_k$ this system has a solution, or, equivalently, C_k is a linear combination of the preceding columns of A if and only if $rank(M_k) = rank(M_{k-1})$. Now the first k column of the echelon matrix B is also an echelon form of M_k. Accordingly, $\mathrm{rank}\,(M_2) = \mathrm{rank}\,(M_3) = 2$ and $\mathrm{rank}\,(M_4) = \mathrm{rank}\,(M_5) = \mathrm{rank}\,(M_6) = 3$. Thus, C_3, C_5, C_6 are each a linear combination of the preceding columns of A.

Now to find basis of column space, the fact that the remaining columns C_1, C_2, C_4 are not linear combinations of their respective preceding columns also tells us that they are linearly independent. Thus, they form a basis of the column space of A. That is, basis of $\mathrm{colsp}(A)$: $[1, 2, 3, 1, 2]^T$, $[2, 5, 7, 5, 6]^T$, $[3, 6, 11, 8, 11]^T$ Observe that C_1, C_2, C_4 may also be characterized as those columns of A that contain the pivots in any echelon form of A. (thus, if it is asked to express a column in terms of basis columns, we can go with equation

$$\sum_{i<k \text{ and } C_i \text{ is a part of basis found above}} x_i C_i = C_k$$

then form an augmented matrix and solve to find variables)

From this exercise we have the following two methods to find basis of the subspace spanned by the vectors $u_1, \ldots, u_r$.

Algorithm 1

1. Form the matrix M whose rows are the given vectors.
2. Row reduce M to echelon form.
3. Output the nonzero rows of the echelon matrix.

But sometimes we want to find a basis that only comes from the original given vectors.

Algorithm 2

1. Form the matrix M whose columns are the given vectors.
2. Row reduce M to echelon form.
3. For each column C_k in the echelon matrix without a pivot, delete (cast out) the vector u_k from the list S of given vectors.
4. Output the remaining vectors in S (which correspond to columns with pivots).

   - **Example:** Let W be the subspace of R^5 spanned by the following vectors: $u_1 = (1, 2, 1, 3, 2)$, $u_2 = (1, 3, 3, 5, 3)$, $u_3 = (3, 8, 7, 13, 8)$ $u_4 = (1, 4, 6, 9, 7)$, $u_5 = (5, 13, 13, 25, 19)$

Find a basis of W consisting of the original given vectors, and find $dim(W)$.

$$M = \begin{bmatrix} 1 & 1 & 3 & 1 & 5 \\ 2 & 3 & 8 & 4 & 13 \\ 1 & 3 & 7 & 6 & 13 \\ 3 & 5 & 13 & 9 & 25 \\ 2 & 3 & 8 & 7 & 19 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 3 & 1 & 5 \\ 0 & 1 & 2 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The pivots in the echelon matrix appear in columns C_1, C_2, C_4. Accordingly, we 'cast out' the vectors u_3 and u_5 from the original five vectors. The remaining vectors u_1, u_2, u_4, which correspond to the columns in the echelon matrix with pivots, form a basis of W. Thus, in particular, $dim(W) = 3$

- Extend $\{u_1 = (1, 1, 1, 1), u_2 = (2, 2, 3, 4)\}$ to a basis of R^4.

  First form the matrix with rows u_1 and u_2, and reduce to echelon form:

  $$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

  Then $w_1 = (1, 1, 1, 1)$ and $w_2 = (0, 0, 1, 2)$ span the same set of vectors as spanned by u_1 and u_2. Let $u_3 = (0, 1, 0, 0)$ and $u_4 = (0, 0, 0, 1)$. Then w_1, u_3, w_2, u_4 form a matrix in echelon form. Thus, they are linearly independent, and they form a basis of R^4. Hence, u_1, u_2, u_3, u_4 also form a basis of R^4.

- Find Basis of subspace W of R^3 where $W = \{(a, b, c) : a + b + c = 0\}$ i.e. vectors of the form $(x, y, -x-y)$, clearly $(1, 0, -1)$ and $(0, 1, -1)$ spans W and are linearly independent and hence they form the basis.

Theorem: The dimension of the solution space W of a homogeneous system $AX = 0$ is $n - r$, where n is the number of unknowns and r is the rank of the coefficient matrix A.

In the case where the system $AX = 0$ is in echelon form, it has precisely $n - r$ free variables, say $x_{i_1}, x_{i_2}, ..., x_{i_{n-r}}$. Let v_j be the solution obtained by setting $x_{i_j} = 1$ (or any nonzero constant) and the remaining free variables equal to 0. Its straightforward to see that the solutions $v_1, v_2, ..., v_{n-r}$ are linearly independent; hence, they form a basis (as they as well span all the solutions) of the solution space W.

- Find the dimension and a basis of the solution space W of each homogeneous system:

  1. $x + 2y + 2z - s + 3t = 0$

$$x + 2y + 3z + s + t = 0$$

$$3x + 6y + 8z + s + 5t = 0$$

Form an augmented matrix, to get $\begin{bmatrix} 1 & 2 & -1 & 3 & 0 \\ 1 & 2 & 3 & 1 & 1 & 0 \\ 3 & 6 & 8 & 1 & 5 & 0 \end{bmatrix}$ $\sim$

$$\begin{bmatrix} 1 & 2 & 2 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The system in echelon form has two (nonzero) equations in five unknowns. Hence, the system has $5 - 2 = 3$ free variables, which are y, s, t. Thus, $\dim W = 3$. We obtain a basis for W:

1. Set $y = 1, s = 0, t = 0$ to obtain the solution $v_1 = (-2, 1, 0, 0, 0)$
2. Set $y = 0, s = 1, t = 0$ to obtain the solution $v_2 = (5, 0, -2, 1, 0)$
3. Set $y = 0, s = 0, t = 1$ to obtain the solution $v_3 = (-7, 0, 2, 0, 1)$

The set $\{v_1, v_2, v_3\}$ is a basis of the solution space W

$$x + y + 2z = 0$$
2. $$2x + 3y + 3z = 0$$
$$x + 3y + 5z = 0$$

Reduce the coefficient matrix A to echelon form:

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \\ 1 & 3 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 2 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 5 \end{bmatrix}$$

This corresponds to a triangular system with no free variables. Thus, 0 is the only solution; that is, $W = \{0\}$. Hence, $\dim W = 0$.

- Find a homogeneous system whose solution set W is spanned by $\{u_1, u_2, u_3\} = \{(1, -2, 0, 3), \quad (1, -1, -1, 4), \quad (1, 0, -2, 5)\}$

Let $v = (x, y, z, t)$. Then $v \in W$ if and only if v is a linear combination of the vectors u_1, u_2, u_3 that span W. Thus, form the matrix M whose first columns are u_1, u_2, u_3 and whose last column is v, and then row reduce M to echelon form. This yields

$$M = \begin{bmatrix} 1 & 1 & 1 & x \\ -2 & -1 & 0 & y \\ 0 & -1 & -2 & z \\ 3 & 4 & 5 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & x \\ 0 & 1 & 2 & 2x + y \\ 0 & -1 & -2 & z \\ 0 & 1 & 2 & -3x + t \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & x \\ 0 & 1 & 2 & 2x + y \\ 0 & 0 & 0 & 2x + y + z \\ 0 & 0 & 0 & -5x - y + t \end{bmatrix}$$

Then v is a linear combination of u_1, u_2, u_3 if $rank(M) = rank(A)$, where A is the submatrix without column Av. Thus, set the last two entries

14

in the fourth column on the right equal to zero to obtain the required homogeneous system:

$2x + y + z = 0$

$5x + y \quad - t = 0$

Inverse of the matrix by Gauss-Jordan Method

Task is to find A^{-1}. I.e. to find X s.t. $AX = I$. Make an augmented matrix $\tilde{A} = [A\ I]$ Now reduce it Reduced Row Echelon form to get $[I\ H]$ since this corresponds to $IX = H$, thus $X = H$.

From this analysis we observe that A^{-1} is simply the result of elementary row operations and thus is product of elementary matrices therefore A is as well the product of elementary matrices.

Examples:

- Find two non-singular matrices P and Q s.t. PAQ is in the normal (canonical) form (i.e. of the form $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$) where $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 3 & 1 & 1 \end{bmatrix}$

  For series of row operations $r_1, r_2, ...$ applied, apply them to I to get R, similarly for any column operation applied in between, apply it to another I to get C, now $P = R, Q = C$. Obv. P and Q aren't unique.

  :::tip Info It is observed that authors club row and column equivalent together to define **equivalent** (which is of course an equivalence relation) matrices which are obtained from each other via combination of elementary row and column operations. Thus, above example shows that any matrix is equivalent to this **normal** form.

  From above it follows that if A and B are of same order and of same rank then they are equivalent to each other. :::

- Is the matrix $\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \\ 2 & 1 & -3 \end{bmatrix}$ equivalent to I_3

  Here determinant of given matrix is non zero and hence it's rank is 3 and therefore it is equivalent to I_3.

- Express $\begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$ as a product of elementary matrices.

  Question is valid iff $|A| \neq 0$ which is true for the given case.

Let A be reduced into I by elementary row transformations $r_1, r_2, \ldots, r_k \to$
$(r_k) \ldots (r_1) A = I$

$\to A = (r_k \ldots r_1)^{-1} I = r_1^{-1} \ldots r_k^{-1} I$

For the given case, $r_1 = E_{21}(-1), r_2 = E_{31}(-1), r_3 = E_{12}(-3), r_4 = E_{13}(-3)$

And thus answer is $E_{21}(1) E_{31}(1) E_{12}(3) E_{13}(3)$

Rank of product of matrices

Rank of product of matrices cannot exceed rank of either of them.

Proof: Proof is simple, suppose we have two matrices A and B each of rank say p, q resp. then it's possible to transform A into $[G\ O]$ where G has only p non zero rows at top, by left multiplication by P. Similarly it is possible to transform B into H where H has only q non zero columns at left, by right multiplication by Q. Now since P, Q are result of elementary matrices that implies $rank(AB) = rank(PGHQ) = rank(GH)$ which is $\le min(p, q)$

Aliter:

1. RB is a linear combination of the rows of B, where R is a row vector. (easy)
2. The row space of AB is contained in the row space of B. (follows from above)
3. If C is a column vector and AC is defined, then AC is a linear combination of the columns of A. (easy)
4. The column space of AB is contained in the column space of A. (follows from above) [Or $colsp(AB) = rowsp((AB)^T) = rowsp(B^T A^T) = rowsp(A^T) = colsp(A)$]

Cramer's Rule

Let the linear system $Ax = b$ have the same number of equations as the unknowns.

Let $D = |A|$, as discussed before the system $Ax = b$ has unique soln iff $rank(A) = n$ i.e. iff $D \ne 0$ and in this case soln is given by $x_i = \frac{D_i}{D}$ where D_i is the determinant obtained from D by replacing in D the ith column by the column with entries $b_1, \ldots, b_n$.

Hence if system is homogeneous and $D \ne 0$, it has only the trivial soln $x_i = 0$. If $D = 0$, the homogeneous system also has nontrivial solns.

Sums and Direct Sums

Let U and W be subsets of a vector space V. The sum of U and W, written $U + W$, consists of all sums $u + w$ where $u \in U$ and $w \in W$. That is,

$$U + W = \{v : v = u + w, \text{ where } u \in U \text{ and } w \in W\}$$

Now suppose U and W are subspaces of V. Then one can easily show that

1. $U + W$ is a subspace of V.
2. U and W are contained in $U + W$.
3. $U + W$ is the smallest subspace containing U and W; that is, $U + W = span(U, W)$.
4. $W + W = W$

Theorem: Suppose U and W are finite-dimensional subspaces of a vector space V. Then $U + W$ has finite dimension and

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

Proof: Observe that $U \cap W$ is a subspace of both U and W. Suppose $dim(U) = m, dim(W) = n, dim(U \cap W) = r$. Suppose $\{v_1, \dots, v_r\}$ is a basis of $U \cap W$. Thus, we can extend $\{v_i\}$ to a basis of U and to a basis of W; say

$$\{v_1, \dots, v_r, u_1, \dots, u_{m-r}\} \quad \text{and} \quad \{v_1, \dots, v_r, w_1, \dots, w_{n-r}\}$$

are bases of U and W, respectively. Let

$$B = \{v_1, \dots, v_r, u_1, \dots, u_{m-r}, w_1, \dots, w_{n-r}\}$$

Note that B has exactly $m + n - r$ elements. Thus, the theorem is proved if we can show that B is a basis of $U + W$. Because $\{v_i, u_j\}$ spans U and $\{v_i, w_k\}$ spans W, the union $B = \{v_i, u_j, w_k\}$ spans $U + W$. Thus, it suffices to show that B is independent. Suppose,

$$a_1 v_1 + \cdots + a_r v_r + b_1 u_1 + \cdots + b_{m-r} u_{m-r} + c_1 w_1 + \cdots + c_{n-r} w_{n-r} = 0 \quad (1)$$

where a_i, b_j, c_k are scalars. Let

$$v = a_1 v_1 + \cdots + a_r v_r + b_1 u_1 + \cdots + b_{m-r} u_{m-r} \quad (2)$$

By (1) we also have
$$v = -c_1 w_1 - \cdots - c_{n-r} w_{n-r} \quad (3)$$

Because $\{v_i, u_j\} \subseteq U, v \in U$ by (2); and as $\{w_k\} \subseteq W, v \in W$ by (3). Accordingly, $v \in U \cap W$. Now $\{v_i\}$ is a basis of $U \cap W$, and so there exist scalars $d_1, \dots, d_r$ for which $v = d_1 v_1 + \cdots + d_r v_r$. Thus, by (3), we have

$$d_1 v_1 + \cdots + d_r v_r + c_1 w_1 + \cdots + c_{n-r} w_{n-r} = 0$$

But $\{v_i, w_k\}$ is a basis of W, and so is independent. Hence, the above equation forces $c_1 = 0, \dots, c_{n-r} = 0$. Substituting this into (1), we obtain

$$a_1 v_1 + \cdots + a_r v_r + b_1 u_1 + \cdots + b_{m-r} u_{m-r} = 0$$

But $\{v_i, u_j\}$ is a basis of U, and so is independent. Hence, the above equation forces

$$a_1 = 0, \dots, a_r = 0, b_1 = 0, \dots, b_{m-r} = 0$$

Because (1) implies that the a_i, b_j, c_k are all 0, $B = \{v_i, u_j, w_k\}$ is independent, and the theorem is proved.

Example: Let $V = M_{2,2}$, the vector space of 2×2 matrices. Let U consist of those matrices whose second row is zero, and let W consist of those matrices whose second column is zero. Then

$$U = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \right\}, \quad W = \left\{ \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} \right\} \text{ and}$$

$$U + W = \left\{ \begin{bmatrix} a & b \\ c & 0 \end{bmatrix} \right\}, \quad U \cap W = \left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

$$\dim(U + W) = \dim U + \dim V - \dim(U \cap W) = 2 + 2 - 1 = 3$$

- Consider the following subspaces of R^5

$$U = \operatorname{span}(u_1, u_2, u_3) = \operatorname{span}\{(1, 3, -2, 2, 3), \quad (1, 4, -3, 4, 2), \quad (2, 3, -1, -2, 9)\}$$

$$W = \operatorname{span}(w_1, w_2, w_3) = \operatorname{span}\{(1, 3, 0, 2, 1), \quad (1, 5, -6, 6, 3), \quad (2, 5, 3, 2, 1)\}$$

Find a basis and the dimension of (a) $U + W$, (b) $U \cap W$

(a) $U + W$ is the space spanned by all six vectors. Hence, form the matrix whose rows are the given six vectors, and then row reduce to echelon form:

$$\begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 1 & 4 & -3 & 4 & 2 \\ 2 & 3 & -1 & -2 & 9 \\ 1 & 3 & 0 & 2 & 1 \\ 1 & 5 & -6 & 6 & 3 \\ 2 & 5 & 3 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & -3 & 3 & -6 & 3 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 2 & -4 & 4 & 0 \\ 0 & -1 & 7 & -2 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -2 & 2 & 3 \\ 0 & 1 & -1 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The following three nonzero rows of the echelon matrix form a basis of $U + W$:

$$(1, 3, -2, 2, 2, 3), \quad (0, 1, -1, 2, -1), \quad (0, 0, 1, 0, -1)$$

Thus, $\dim(U + W) = 3$

18

(b) Let $v = (x, y, z, s, t)$ denote an arbitrary element in R^5. First find, like before, homogeneous systems whose solution sets are U and W, respectively. Let M be the matrix whose columns are the u_i and v, and reduce M to echelon form:

$$M = \begin{bmatrix} 1 & 1 & 2 & x \\ 3 & 4 & 3 & y \\ -2 & -3 & -1 & z \\ 2 & 4 & -2 & s \\ 3 & 2 & 9 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & x \\ 0 & 1 & -3 & -3x + y \\ 0 & 0 & 0 & -x + y + z \\ 0 & 0 & 0 & 4x - 2y + s \\ 0 & 0 & 0 & -6x + y + t \end{bmatrix}$$

That implies:

$$-x + y + z = 0, \quad 4x - 2y + s = 0, \quad -6x + y + t = 0$$

Similarly,

$$M' = \begin{bmatrix} 1 & 1 & 2 & x \\ 3 & 5 & 5 & y \\ 0 & -6 & 3 & z \\ 2 & 6 & 2 & s \\ 1 & 3 & 1 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & x \\ 0 & 2 & -1 & -3x + y \\ 0 & 0 & 0 & -9x + 3y + z \\ 0 & 0 & 0 & 4x - 2y + s \\ 0 & 0 & 0 & 2x - y + t \end{bmatrix}$$

That implies:

$$-9 + 3 + z = 0, \quad 4x - 2y + s = 0, \quad 2x - y + t = 0$$

Combine both of the above systems to obtain a homogeneous system, whose solution space is $U \cap W$, and reduce the system to echelon form, yielding

$$-x + y + z = 0 \quad 2y + 4z + s = 0 \quad 8z + 5s + 2t = 0 \quad s - 2t = 0$$

There is one free variable, which is t; hence, $dim(U \cap W) = 1$. Setting $t = 2$, we obtain the solution $u = (1, 4, -3, 4, 2)$, which forms our required basis of $U \cap W$

:::caution IMP Note Here we can't just reduce $A = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ to row canonical

matrix R_1 and $B = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ to row canonical matrix R_2 and take common

rows in two to find basis of $U \cap W$ as common elements can be linear combination of the rows in R_1! :::

- Suppose U and W are distinct four-dimensional subspaces of a vector space V, where dim $V = 6$. Find the possible dimensions of $U \cap W$.

Because U and W are distinct, $U + W$ properly contains U and W; consequently, $dim(U + W) > 4$. But $dim(U + W)$ cannot be greater than 6, as dim $V = 6$. Hence, we have two possibilities:

1. $dim(U + W) = 5$ and thus $dim(U \cap W) = 3$
2. $dim(U + W) = 6$ and thus $dim(U \cap W) = 2$

Direct Sums

The vector space V is said to be the direct sum of its subspaces U and W, denoted by

$$V = U \oplus W$$

if every $v \in V$ can be written in one and only one way as $v = u + w$ where $u \in U$ and $w \in W$.

Theorem: The vector space V is the direct sum of its subspaces U and W if and only if:

(i) $V = U + W$, (ii) $U \cap W = \{0\}$ (easy to prove)

General Direct Sums

V is the direct sum of subspaces $W_1, W_2, \dots, W_r$, written $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$

if every vector $v \in V$ can be written in one and only one way as

$$v = w_1 + w_2 + \cdots + w_r$$

where $w_1 \in W_1, w_2 \in W_2, \dots, w_r \in W_r$

Theorem: Suppose $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$. Also, for each k, suppose S_k is a linearly independent subset of W_k. Then

(a) The union $S = \cup_k S_k$ is linearly independent in V (easy to prove as, if they are linearly dependent then one side is a linear combination of S_1 (which would imply coefficients on left side are 0 and so on...))

(b) If each S_k is a basis of W_k, then $\bigcup_k S_k$ is a basis of V (follows from (a))

(c) $\dim V = \dim W_1 + \dim W_2 + \cdots + \dim W_r$ (follows from (b))

Theorem: Suppose $V = W_1 + W_2 + \cdots + W_r$ and $\dim V = \sum_k \dim W_k$. Then

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$$

Coordinates

Let V be an n-dimensional vector space over K with basis $S = \{u_1, u_2, \dots, u_n\}$. Then any vector $v \in V$ can be expressed uniquely as a linear combination of the basis vectors in S (this is a theorem for basis which is easy to prove), say $v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$ These n scalars $a_1, a_2, \dots, a_n$ are called the

coordinates of v relative to the basis S, and they form a vector $[a_1, a_2, \ldots, a_n]$ in K^n called the coordinate vector of v relative to S. We denote this vector by $[v]_S$, or simply $[v]$, when S is understood.

- In the space $M = M_{2,3}$, determine whether or not the following matrices are linearly dependent:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 & 7 \\ 10 & 1 & 13 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 & 5 \\ 8 & 2 & 11 \end{bmatrix}$$

If the matrices are linearly dependent, find the dimension and a basis of the subspace W of M spanned by the matrices

The coordinate vectors of the above matrices relative to the usual basis of M are as follows:

$$[A] = [1, 2, 3, 4, 0, 5], \quad [B] = [2, 4, 7, 10, 1, 13], \quad [C] = [1, 2, 5, 8, 2, 11]$$

$$M = \begin{bmatrix} 1 & 2 & 3 & 4 & 0 & 5 \\ 2 & 4 & 7 & 10 & 1 & 13 \\ 1 & 2 & 5 & 8 & 2 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 4 & 0 & 5 \\ 0 & 0 & 1 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Because the echelon matrix has only two nonzero rows, the coordinate vectors $[A], [B], [C]$ span a space of dimension two, and so they are linearly dependent. Thus, A, B, C are linearly dependent. Furthermore, $\dim W = 2$, and the matrices

$$w_1 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 5 \end{bmatrix} \quad \text{and} \quad w_2 = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

corresponding to the nonzero rows of the echelon matrix form a basis of W.

Linear Mappings

Matrix Mappings

Let A be any $m \times n$ matrix over K. Then A determines a mapping $F_A : K^n \to K^m$ by

$$F_A(u) = Au$$

where the vectors in K_n and K_m are written as columns.

For notational convenience, we will frequently denote the mapping F_A by the letter A, the same symbol as used for the matrix.

Linear Mappings (Linear Transformations)

Let V and U be vector spaces over the same field K. A mapping $F : V \to U$ is called a linear mapping or linear transformation if it satisfies the following two conditions:

1. For any vectors $v, w \in V$, $F(v + w) = F(v) + F(w)$
2. For any scalar k and vector $v \in V$, $F(kv) = kF(v)$

Substituting $k = 0$ into condition (2), we obtain $F(0) = 0$. Thus, every linear mapping takes the zero vector into the zero vector.

For any scalars $a_i \in K$ and any vectors $v_i \in V$, we obtain the following basic property of linear mappings:

$$F\left(a_1 v_1 + a_2 v_2 + \cdots + a_m v_m\right) = a_1 F\left(v_1\right) + a_2 F\left(v_2\right) + \cdots + a_m F\left(v_m\right)$$

Remark: A linear mapping $F : V \to U$ is completely characterized by the condition

$$F(av + bw) = aF(v) + bF(w)$$

1

Theorem 4.1: Let V and U be vector spaces over a field K. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of V and let $u_1, u_2, \ldots, u_n$ be any vectors in U. Then there exists a unique linear mapping $F : V \to U$ such that $F(v_1) = u_1, F(v_2) = u_2, \ldots, F(v_n) = u_n$.

Note that the vectors $u_1, u_2, \ldots, u_n$ are completely arbitrary; they may be linearly dependent or they may even be equal to each other.

Proof: Consider $v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$

Define $F(v) = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$

Suppose $G : V \to U$ is linear and $G(v_1) = u_i, i = 1, \ldots, n$. Let

Then

$G(v) = G(a_1 v_1 + a_2 v_2 + \cdots + a_n v_n) = a_1 G(v_1) + a_2 G(v_2) + \cdots + a_n G(v_n)$

$= a_1 u_1 + a_2 u_2 + \cdots + a_n u_n = F(v)$

Because $G(v) = F(v)$ for every $v \in V, G = F$. Thus, F is unique and the theorem is proved.

Matrices as Linear Mappings

$F_A(v + w) = A(v + w) = Av + Aw = F_A(v) + F_A(w)$
$\quad F_A(kv) = A(kv) = k(Av) = kF_A(v)$

In other words, using A to represent the mapping, we have,

$A(v + w) = Av + Aw \quad$ and $\quad A(kv) = k(Av)$

Vector Space Isomorphism

Two vector spaces V and U over K are isomorphic, written $V \cong U$, if there exists a bijective (one-to-one and onto) linear mapping $F : V \to U$. The mapping F is then called an isomorphism between V and U.

Consider any vector space V of dimension n and let S be any basis of V. Then the mapping

$v \mapsto [v]_S$

which maps each vector $v \in V$ into its coordinate vector $[v]_S$, is an isomorphism between V and K^n

Kernal and Image of a Linear Mapping

Let $F : V \to U$ be a linear mapping. The kernel of F, written $\operatorname{Ker} F$, is:

$\operatorname{Ker} F = \{v \in V : F(v) = 0\}$

And image or range of F is defined as:

$\operatorname{Im} F = \{u \in U : \text{ there exists } v \in V \text{ for which } F(v) = u\}$

Theorem 4.2 Let $F : V \to U$ be a linear mapping. Then the kernel of F is a subspace of V and the image of F is a subspace of U. (easy to see)

Theorem 4.3 Suppose $v_1, v_2, \ldots, v_m$ span a vector space V, and suppose $F : V \to U$ is linear. Then $F(v_1), F(v_2), \ldots, F(v_m)$ span $\operatorname{Im} F$. (easy to see)

Thus one can either use this, or theorem 4.4 to find the dimension of $\operatorname{Im} F$

Examples:

- Let $F : R^3 \to R^3$ be the projection of a vector v into the xy-plane that is, $F(x, y, z) = (x, y, 0)$ Clearly the image of F is the entire xy-plane—that is, points of the form $(x, y, 0)$. Moreover, the kernel of F is the z-axis—that is, points of the form $(0, 0, c)$.

Kernal and Image of Matrix Mappings

Let A by any $m \times n$ matrix over a field K viewed as a linear map $A : K^n \times K^m$. Then

1. $\operatorname{Im} A = \operatorname{colsp}(A)$ as consider usual basis e_i of K^n, $Ae_1, Ae_2, \ldots, Ae^n$ are respectively the columns of A.
2. Kernel of A consists of all vectors v for which $Av = 0$. This means that the kernel of A is the solution space of the homogeneous system $AX = 0$, called the null space of A.

If we have computed the dimension of column space, which is same as rank of the matrix, r. And hence the dimension of kernal of A is $n - r$.

Examples:

- Find a linear map $F : R^3 \to R^4$ whose image is spanned by $(1, 2, 0, -4)$ and $(2, 0, 1, -3)$.

  $u = (x + 2y, 2x, -y, -4x - 3y)$

$$\therefore A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 0 & 0 \\ 0 & -1 & -1 \\ -4 & -3 & -3 \end{bmatrix}$$

Also recall that A determines a linear map $A : R^3 \to R^4$ whose image is spanned by the columns of A. Thus, A satisfies the required condition

Rank and Nullity of a Linear Mapping

Let $F : V \to U$ be a linear mapping.

$\text{rank}(F) = \dim(\text{Im } F)$ and nullity $(F) = \dim(\text{Ker } F)$

Theorem 4.4 Let V be of finite dimension, and let $F : V \to U$ be linear. Then

$\dim V = \dim(\text{Ker } F) + \dim(\text{Im } F) = $ nullity $(F) + \text{rank}(F)$

Singular and Nonsingular Linear Mappings, Isomorphisms

Let $F : V \to U$ be a linear mapping. Recall that $F(0) = 0$. F is said to be singular if the image of some nonzero vector v is 0. Thus, $F : V \to U$ is nonsingular if the zero vector 0 is the only vector whose image under F is 0 or, in other words, if $\text{Ker } F = \{0\}$

_____

Theorem 4.5: Let $F : V \to U$ be a nonsingular linear mapping. Then the image of any linearly independent set is linearly independent.

_____

_____

Theorem 4.6: A linear mapping $F : V \to U$ is one-to-one if and only if F is nonsingular. (easy to see)

_____

_____

Theorem 4.7: Suppose V has finite dimension and $dimV = dimU$. Suppose $F : V \to U$ is linear. Then F is an isomorphism if and only if F is nonsingular.

_____

4

Operations with Linear Mappings

Let $F : V \to U$ and $G : V \to U$ be linear mappings over a field K. The sum $F + G$ and the scalar product kF, where $k \in K$, are defined to be the following mappings from V into U:

$$(F + G)(v) \equiv F(v) + G(v) \quad \text{and} \quad (kF)(v) \equiv kF(v)$$

It is easy to see that if F and G are linear, then $F + G$ and kF are also linear.

Theorem 4.8: Let V and U be vector spaces over a field K. Then the collection of all linear mappings from V into U with the above operations of addition and scalar multiplication forms a vector space over K.

TODO

Left 5.21